AD-763 461

ESTIMATING AND TESTING LINEAR HYPOTHESES ON PARAMETERS IN THE LOG-LINEAR MODEL

S. Kullback

George Washington University

Prepared for:

Office of Naval Research

28 May 1973

Unclassified

Security Classification

AD 763461

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) THE GEORGE WASHINGTON UNIVERSITY DEPARTMENT OF STATISTICS WASHINGTON, D.C. 20006	2a. REPORT SECURITY CLASSIFICATION 2b. GROUP

3. REPORT TITLE

ESTIMATING AND TESTING LINEAR HYPOTHESES ON PARAMETERS IN THE LOG-LINEAR MODEL

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

TECHNICAL REPORT

5. AUTHOR(S) (First name, middle initial, last name)

S. KULLBACK

6. REPORT DATE	7a. TOTAL NO OF PAGES	7b. NO OF REFS
May 28, 1973	~~33~~ 36	15

8a. CONTRACT OR GRANT NO N00014-67-A-0214-0015 b. PROJECT NO NR-042-267 c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) #6 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. DISTRIBUTION STATEMENT

Unlimited. Reproduction in whole or in part is permitted for any purpose of the United States Government.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	OFFICE OF NAVAL RESEARCH STATISTICS & PROBABILITY PROGRAM ARLINGTON, VIRGINIA 22217

13. ABSTRACT

The minimum discrimination information theorem yields minimum discrimination information estimates as members of an exponential family which can be expressed in log-linear additive form. We illustrate in terms of a particular contingency table which has been examined from other points of view the relations implied by hypotheses and subhypotheses of interest on the values of the natural parameters and random variables of the exponential family, the related moment parameters, and the estimates of cell entries. Since the estimates are constrained to satisfy certain linear relations based on observed values, the minimum discrimination information estimates are maximum-likelihood estimates and the associated minimum discrimination information statistics are log-likelihood ratio statistics. Computations were carried out using a general computer program for the analysis of contingency tables including the Deming-Stephan iterative algorithm and its extension by the generalized iterative scaling procedure of Darroch and Ratcliff. Results are summarized in Analysis of Information tables.

Unclassified
Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
CONTINGENCY TABLE ANALYSIS						
MINIMUM DISCRIMINATION INFORMATION						
ANALYSIS OF INFORMATION						
LOG-LINEAR MODEL						
EXPONENTIAL FAMILY						

DD FORM 1 NOV 65 1473 (BACK)
(PAGE 2)

Unclassified
Security Classification

ESTIMATING AND TESTING LINEAR HYPOTHESES ON

PARAMETERS IN THE LOG-LINEAR MODEL

by

S. KULLBACK

TECHNICAL REPORT NO. 6

May 28, 1973

PREPARED UNDER CONTRACT N00014-67-A-0214-0015

(NR-042-267)

OFFICE OF NAVAL RESEARCH

Herbert Solomon, Project Director

DEPARTMENT OF STATISTICS

THE GEORGE WASHINGTON UNIVERSITY

WASHINGTON, D.C. 20006

Estimating and Testing Linear Hypotheses on Parameters in the Log-linear Model

by

S. Kullback

The George Washington University

Summary

The minimum discrimination information theorem yields minimum discrimination information estimates as members of an exponential family which can be expressed in log-linear additive form. We illustrate in terms of a particular contingency table the relations implied by hypotheses and subhypotheses of interest on the values of the natural parameters and random variables of the exponential family, the related moment parameters, and the estimates of cell entries. Computations are carried out using the Deming-Stephan iterative algorithm and its extension by the generalised iterative scaling procedure of Darroch and Ratcliff.

1. Introduction

It is a consequence of the minimum discrimination information theorem, that minimum discrimination information distributions are formulated as members of an exponential family [3], [11]. For applications to the analysis of contingency tables it is useful to express the exponential

family as a log-linear additive model [10], [13]. We propose to illustrate in terms of a particular set of data the relations implied by hypotheses and subhypotheses of interest on the values of the natural parameters and random variables of the exponential family, the related moment parameters, and the estimates of the cell entries.

Since our estimates in this discussion are constrained to satisfy certain linear relations based on observed values, the minimum discrimination information estimates are maximum likelihood estimates and the associated minimum discrimination information statistics are log-likelihood ratio statistics [3], [5], [12], [13]. Further references to the literature on the log-linear model may be found in papers by Bishop [2], Dempster [6], Gokhale [7], Ku et. al. [10], Plackett [14].

The particular data we shall use were first presented by Cochran [4], are given in table 1, and have been examined from various points of view [1], [2], [7], [8], [10], [15]. It should be noted that the use of the data is primarily as a vehicle to illustrate the procedures and concepts rather than the reanalysis of the data per se. Computations were carried out using a general computer program for the analysis of contingency tables including the Deming-Stephan iterative algorithm and its extension by the generalised iterative scaling procedure of Darroch and Ratcliff [5].

2. The Hypothesis and Subhypotheses

As stated by Grizzle [8, p. 379], "It is desired to compare the mothers of Baltimore school children who have been referred by their teachers as presenting behavioral problems to mothers of a comparable group of control children. For each mother it is recorded whether she had suffered any infant losses previous to the child in the study." The hypotheses considered by Grizzle may be stated as [1], [8]:

Case 1. "The test of the hypothesis of no interaction" or logit additivity.

Then "to test hypotheses about the parameters in the model."

Case 2. "Is there a difference between problems and controls?"

Case 3. Is there "equality of birth-order effects?"

Case 4. a) "Is the effect of birth-order linear?"

b) "Does it require a second-degree polynomial to describe the effect?"

It was shown that the null hypothesis of logit additivity of case 1 is satisfied [8]. Maximum-likelihood estimates for the subhypotheses of cases 2,3,4 were obtained in terms of linear contrasts of the logits, but did not include a specific restraint of logit additivity [8]. Berkson [1] gave the maximum-likelihood estimates, as well as minimum logit χ^2 estimates under the same

restraints of the linear contrasts of the logits. The maximum-likelihood estimates are given in table 2. The resulting maximum-likelihood estimates for cases 2,3,4 consequently do not satisfy logit additivity as shown in (2.1)-(2.4). Similar results also follow for the minimum logit χ^2 estimates. The procedure we shall present derives maximum-likelihood estimates for the various cases that do satisfy logit additivity.

3. The log-linear Representation

In accordance with the minimum discrimination information theorem [11, p. 38], [13] a log-linear representation for the observed values x(ijk) of table 1 for the complete set of marginals, including observed cell values (as three-way marginals), is given analytically by (3.1) where $\pi(ijk)$ is the uniform distribution and n is the total number of observations, and graphically in table 3a. The values in column 1 of table 3a are essentially those of a normalizing constant (the negative of the logarithm of the moment-generating function). The values in the other columns of table 3a are those of a set of linearly independent functions T(ijk) whose mean values are various marginals. Note that

$$T_{11}^{ij}(ijk) = T_1^i(ijk)T_1^j(ijk), \text{etc.}; T_{111}^{ijk}(ijk)$$
$$= T_1^i(ijk)T_1^j(ijk)T_1^k(ijk), \text{etc},$$

where the T(ijk) are indicator functions of the various marginals [13].

$$\text{Case 2}\quad\left\{\begin{aligned}&\ln\frac{17.812}{84.188}-\ln\frac{12.188}{51.182}=\ln 0.980\\&\ln\frac{23.812}{43.188}-\ln\frac{18.188}{27.812}=\ln 0.843\\&\ln\frac{24.812}{24.188}-\ln\frac{16.188}{20.812}=\ln 1.319\end{aligned}\right.\tag{2.1}$$

$$\text{Case 3}\quad\left\{\begin{aligned}&\ln\frac{29.925}{72.075}-\ln\frac{19.925}{44.075}=\ln 0.918\\&\ln\frac{22.471}{44.529}-\ln\frac{12.471}{33.529}=\ln 1.357\\&\ln\frac{20.604}{28.396}-\ln\frac{7.604}{29.396}=\ln 2.805\end{aligned}\right.\tag{2.2}$$

$$\text{Case 4a}\quad\left\{\begin{aligned}&\ln\frac{27.535}{74.465}-\ln\frac{17.535}{46.465}=\ln 0.980\\&\ln\frac{26}{41}-\ln\frac{16}{30}=\ln 1.189\\&\ln\frac{19.465}{29.535}-\ln\frac{6.465}{30.535}=\ln 3.113\end{aligned}\right.\tag{2.3}$$

$$\text{Case 4b}\quad\left\{\begin{aligned}&\ln\frac{21.191}{30.809}-\ln\frac{11.191}{52.809}=\ln 1.237\\&\ln\frac{23.618}{43.382}-\ln\frac{13.618}{32.382}=\ln 1.295\\&\ln\frac{28.191}{20.809}-\ln\frac{15.191}{21.809}=\ln 1.945\end{aligned}\right.\tag{2.4}$$

$$\begin{aligned}\ln\frac{x(ijk)}{n\pi(ijk)} &= L + \tau_1^i T_1^i(ijk) + \tau_1^j T_1^j(ijk)\\&+ \tau_1^k T_1^k(ijk) + \tau_2^k T_2^k(ijk) + \tau_{11}^{ij} T_{11}^{ij}(ijk)\\&+ \tau_{11}^{ik} T_{11}^{ik}(ijk) + \tau_{12}^{ik} T_{12}^{ik}(ijk) + \tau_{11}^{jk} T_{11}^{jk}(ijk)\\&+ \tau_{12}^{jk} T_{12}^{jk}(ijk) + \tau_{111}^{ijk} T_{111}^{ijk}(ijk) + \tau_{112}^{ijk} T_{112}^{ijk}(ijk)\end{aligned}\tag{3.1}$$

From table 3a and the representation in (3.1) we may express the logits over losses/no losses in the observed contingency table parametrically as in (3.2). The values of the effect and interaction (tau) parameters for x(ijk) are given in table 5. Let us now consider the implication of the hypotheses.

$$
\begin{aligned}
\ln\frac{x(111)}{x(121)} &= \tau_{1}^{j} + \tau_{11}^{ij} + \tau_{11}^{jk} + \tau_{111}^{ijk} \\
\ln\frac{x(112)}{x(122)} &= \tau_{1}^{j} + \tau_{11}^{ij} + \tau_{12}^{jk} + \tau_{112}^{ijk} \\
\ln\frac{x(113)}{x(123)} &= \tau_{1}^{j} + \tau_{11}^{ij} \\
\ln\frac{x(211)}{x(221)} &= \tau_{1}^{j} + \tau_{11}^{jk} \\
\ln\frac{x(212)}{x(222)} &= \tau_{1}^{j} + \tau_{12}^{jk} \\
\ln\frac{x(213)}{x(223)} &= \tau_{1}^{j}
\end{aligned}
\qquad (3.2)
$$

Case 1

For logit additivity (no second-order interaction) the estimate must satisfy $\tau_{111}^{ijk}=\tau_{112}^{ijk}=0$, in order that the difference between corresponding logits for problems and controls (i=1,2) be the same. Its representation is given by columns 1-10 of table 3a. The minimum discrimination information estimates are thus constrained to have the same two-way marginals as the observed table (implying all lower order marginals) and may be obtained by

the Deming-Stephan iterative algorithm. This procedure was illustrated for this data by Ku et al [10, p. 59]. If we denote the estimates by fitting the two-way marginals, that is, satisfying logit additivity, as $x_2^*(ijk)$, then the corresponding logits may be expressed parametrically as in (3.3).

$$
\begin{aligned}
\ln\frac{x_2^*(111)}{x_2^*(121)} &= \tau_1^j + \tau_{11}^{ij} + \tau_{11}^{jk} &&= \mu + \alpha_1 + \beta_1 \\
\ln\frac{x_2^*(112)}{x_2^*(122)} &= \tau_1^j + \tau_{11}^{ij} + \tau_{12}^{jk} &&= \mu + \alpha_1 + \beta_2 \\
\ln\frac{x_2^*(113)}{x_2^*(123)} &= \tau_1^j + \tau_{11}^{ij} &&= \mu + \alpha_1 + \beta_3 \\
\ln\frac{x_2^*(211)}{x_2^*(221)} &= \tau_1^j + \tau_{11}^{jk} &&= \mu + \alpha_2 + \beta_1 \\
\ln\frac{x_2^*(212)}{x_2^*(222)} &= \tau_1^j + \tau_{12}^{jk} &&= \mu + \alpha_2 + \beta_2 \\
\ln\frac{x_2^*(213)}{x_2^*(223)} &= \tau_1^j &&= \mu + \alpha_2 + \beta_3
\end{aligned}
\tag{3.3}
$$

We remark that the numerical values of the tau parameters in (3.3) are of course not the same as those in (3.2), even though we use the same notation for simplicity. We may also consider a parametrization in terms of orthogonal polynomials [9, p. 359] as in (3.4) assuming equal spacing of birth-order effects [8].

Grizzle used the parametrization of (3.3) in μ, α_i, β_j with $\alpha_1+\alpha_2=0$, $\beta_1+\beta_2+\beta_3=0$ [8].

$$
\begin{aligned}
&\tau_1^j + \tau_{11}^{ij} + \tau_{11}^{jk} &&= a_1 - b_1 + b_2 \\
&\tau_1^j + \tau_{11}^{ij} + \tau_{12}^{jk} &&= a_1 - 2b_2 \\
&\tau_1^j + \tau_{11}^{ij} &&= a_1 + b_1 + b_2 \\
&\tau_1^j + \tau_{11}^{jk} &&= a_2 - b_1 + b_2 \\
&\tau_1^j + \tau_{12}^{jk} &&= a_2 - 2b_2 \\
&\tau_1^j &&= a_2 + b_1 + b_2
\end{aligned}
\tag{3.4}
$$

To test the null hypothesis of case 1 we compute the statistic

$$2I(x:x_2^*) = 2\sum\sum\sum x(ijk)\ \ell n(x(ijk)/x_2^*(ijk)) = 0.853,\ 2D.F.$$

The null hypothesis of logit additivity is accepted. Note that this is denoted as Model (1) by Bishop [2, p. 556].

We see from (3.3) and (3.4) that

$$
\begin{aligned}
&\tau_1^j = \mu + \alpha_2 + \beta_3 = a_2 + b_1 + b_2 \\
&\tau_{11}^{ij} = \alpha_1 - \alpha_2 = a_1 - a_2 \\
&\tau_{11}^{jk} = \beta_1 - \beta_3 = -2b_1 \\
&\tau_{12}^{jk} \quad - \beta_3 = -b_1 - 3b_2 \\
&\tau_{11}^{jk} - 2\tau_{12}^{jk} = \beta_1 - 2\beta_2 + \beta_3 = 6b_2
\end{aligned}
\tag{3.5}
$$

The values of $x_2^*(ijk)$ are given in table 4 and the common logit difference is τ_{11}^{ij}. The values of the parameters for x_2^* are given in table 5.

Case 2

For the subhypothesis that there is no difference in the logits between problems i=1 and controls i=2 in the model of logit additivity we must have

$$\tau_{11}^{ij} = \alpha_1 - \alpha_2 = a_1 - a_2 = 0$$

in (3.3)-(3.5), that is, the corresponding estimate $x_e^*(ijk)$ must satisfy

$$\tau_{111}^{ijk} = \tau_{112}^{ijk} = \tau_{11}^{ij} = 0.$$

The representation for $x_e^*(ijk)$ is given by columns 1-5, 7-10 of table 3a. The estimate $x_e^*(ijk)$ must thus have the two-way marginals $x_e^*(i.k) = x(i.k)$, $x_e^*(.jk) = x(.jk)$ and the implied lower order marginals. In this case the estimate is explicitly given by

$$x_e^*(ijk) = x(i.k)x(.jk)/x(..k).$$

Note that this is denoted as PA Model (2) by Bishop [2, p. 556]. The numerical values of $x_e^*(ijk)$ are given in table 4 and the common logit difference in this case is zero. The parameter values for x_e^* are given in table 5.

Case 3

For the subhypothesis that there is equality of birth-order effects k=1,2,3 in the model of logit additivity we must have $\beta_1=\beta_2=\beta_3$ or $\tau_{11}^{jk}=\tau_{12}^{jk}=0$ or $b_1=b_2=0$ in (3.3)-(3.5), that is, the corresponding estimate $x_f^*(ijk)$ must satisfy

$$\tau_{111}^{ijk} = \tau_{112}^{ijk} = \tau_{11}^{jk} = \tau_{12}^{jk} = 0 \ .$$

The representation for $x_f^*(ijk)$ is given by columns 1-8 of table 3a. The estimate $x_f^*(ijk)$ must thus have the two-way marginals $x_f^*(ij.)=x(ij.)$, $x_f^*(i.k)=x(i.k)$ and the implied lower order marginals. In this case the estimate is explicitly given by (a version of Bishop's PA model [2,p. 556])

$$x_f^*(ijk) = x(ij.)x(i.k)/x(i..).$$

The numerical values of $x_f^*(ijk)$ are given in table 4. The common logit difference is τ_{11}^{ij}. The parameter values for x_f^* are given in table 5.

Case 4a

For the null subhypothesis Ho_1 that the linear component is zero in the model of logit additivity we must have $\beta_1=\beta_3$ or $\tau_{11}^{jk} = 0$ or $b_1 = 0$ in (3.3)-(3.5), that is, the corresponding estimate $x_g^*(ijk)$

must satisfy $\tau_{111}^{ijk} = \tau_{112}^{ijk} = \tau_{11}^{jk} = 0$. The representation for $x_g^*(ijk)$ is given by columns 1-8,10 of table 3a. The estimate $x_g^*(ijk)$ must thus satisfy the restraints

$$x_g^*(ij.) = x(ij.),\ x_g^*(i.k) = x(i.k),\ x_g^*(.j2) = x(.j2)$$

on the two-way marginals and the implied lower order marginals. The values of $x_g^*(ijk)$ are obtained by the iteration (cf. Gokhale [7])

$$x^{(3n+1)}(ijk) = \frac{x(ij.)}{x^{(3n)}(ij.)}\, x^{(3n)}(ijk)$$

$$x^{(3n+2)}(ijk) = \frac{x(i.k)}{x^{(3n+1)}(i.k)}\, x^{(3n+1)}(ijk)$$

$$x^{(3n+3)}(ijk) = \frac{x(.jk)}{x^{(3n+2)}(.jk)}\, x^{(3n+2)}(ijk),\ k=2 \qquad (3.6)$$

$$x^{(3n+3)}(ijk) = \frac{x(..1)+x(..3)}{x^{(3n+2)}(..1)+x^{(3n+2)}(..3)}\, x^{(3n+2)}(ijk),$$

$k=1,3$.

The numerical values of $x_g^*(ijk)$ are given in table 4. The common logit difference is τ_{11}^{ij}. The parameter values for x_g^* are given in table 5.

Case 4b

For the subhypothesis Ho_2 that the quadratic component is zero in the model of logit additivity, we must have $\beta_1-2\beta_2+\beta_3=0$ or

$2\tau_{12}^{jk}=\tau_{11}^{jk}$ or $b_2=0$ in (3.3)-(3.5), that is, the corresponding estimate $x_h^*(ijk)$ must satisfy $\tau_{111}^{ijk} = \tau_{112}^{ijk} = 0$, $2\tau_{12}^{jk} = \tau_{11}^{jk} = \tau^{jk}$ where we use the parameter τ^{jk} to avoid confusion. To determine the restraints and estimation procedure for $x_h^*(ijk)$ we note that the requirement $2\tau_{12}^{jk} = \tau_{11}^{jk} = \tau^{jk}$ implies that in the log-linear regression (3.1) there appears a term of the form

$$\ldots\tau^{jk}(T_{11}^{jk}(ijk) + \tfrac{1}{2}T_{12}^{jk}(ijk))\ldots \tag{3.7}$$

so that the corresponding revised graphic table for $x_h^*(ijk)$ appears as in table 3b. The estimate $x_h^*(ijk)$ must satisfy the restraints

$$x_h^*(ij.) = x(ij.),\ x_h^*(i.k) = x(i.k)$$

$$x_h^*(.j1) + x_h^*(.j2)/2 = x(.j1) + x(.j2)/2$$

and those implied on the lower order marginals. The values of $x_h^*(ijk)$ are determined by the generalised iterative scaling procedure of Darroch and Ratcliff [5] (see appendix)

$$\begin{aligned} x^{(3n+1)}(ijk) &= (x(ij.)/x^{(3n)}(ij.))x^{(3n)}(ijk) \\ x^{(3n+2)}(ijk) &= (x(i.k)/x^{(3n+1)}(i.k))x^{(3n+1)}(ijk) \\ x^{(3n+3)}(ijk) &= \end{aligned} \tag{3.8}$$

$$\left(\frac{h_1}{h_1^{(3n+2)}}\right)^{a_1(ijk)} \left(\frac{h_2}{h_2^{(3n+2)}}\right)^{a_2(ijk)} \left(\frac{h_3}{h_3^{(3n+2)}}\right)^{a_3(ijk)} x^{(3n+2)}(ijk)$$

where the values for $a_r(ijk)$, h_r, r=1,2,3 are given in table 6. The numerical values of $x_h^*(ijk)$ are given in table 4. The common logit difference is τ_{11}^{ij}. The parameter values for x_h^* are given in table 5.

Remarks

It is of interest to note that the procedure used by Grizzle [8] which leads to the estimates $x_a^*, x_b^*, x_c^*, x_d^*$ for case 2,3,4a,4b respectively, given in table 2, is in fact equivalent to using the parametric tau representation in (3.2) rather than that in (3.3). The contrasts in terms of the logits thus imply the following relations among the tau parameters.

Case 2

Contrast: $$\ln\frac{x(111)}{x(121)} + \ln\frac{x(112)}{x(122)} + \ln\frac{x(113)}{x(123)} = \ln\frac{x(211)}{x(221)} + \ln\frac{x(212)}{x(222)} + \ln\frac{x(213)}{x(223)}$$

$$3\tau_1^j + 3\tau_{11}^{ij} + \tau_{11}^{jk} + \tau_{12}^{jk} + \tau_{111}^{ijk} + \tau_{112}^{ijk} = 3\tau_1^j + \tau_{11}^{jk} + \tau_{12}^{jk},$$

that is,

$$3\tau_{11}^{ij} + \tau_{111}^{ijk} + \tau_{112}^{ijk} = 0 .$$

Case 3

Contrast: $$\ln\frac{x(111)}{x(121)} + \ln\frac{x(211)}{x(221)} = \ln\frac{x(112)}{x(122)} + \ln\frac{x(212)}{x(222)} = \ln\frac{x(113)}{x(123)} + \ln\frac{x(213)}{x(223)}$$

$$2\tau_1^j + \tau_{11}^{ij} + 2\tau_{11}^{jk} + \tau_{111}^{ijk} = 2\tau_1^j + \tau_{11}^{ij} + 2\tau_{12}^{jk} + \tau_{112}^{ijk} = 2\tau_1^j + \tau_{11}^{ij},$$

that is

$$2\tau_{11}^{jk} + \tau_{111}^{ijk} = 0, \quad 2\tau_{12}^{jk} + \tau_{112}^{ijk} = 0.$$

Case 4a

Contrast: $\ell n\frac{x(111)}{x(121)}+\ell n\frac{x(211)}{x(221)}=\ell n\frac{x(113)}{x(123)}+\ell n\frac{x(213)}{x(223)}$

$$2\tau_{1}^{j}+\tau_{11}^{ij}+2\tau_{11}^{jk}+\tau_{111}^{ijk}=2\tau_{1}^{j}+\tau_{11}^{ij},$$

that is,

$$2\tau_{11}^{jk}+\tau_{111}^{ijk}=0.$$

Case 4b

Contrast: $\ell n\frac{x(111)}{x(121)}+\ell n\frac{x(113)}{x(123)}+\ell n\frac{x(211)}{x(221)}+\ell n\frac{x(213)}{x(223)}=2\left(\ell n\frac{x(112)}{x(122)}+\ell n\frac{x(212)}{x(222)}\right)$

$$4\tau_{1}^{j}+2\tau_{11}^{ij}+2\tau_{11}^{jk}+\tau_{111}^{ijk}=2(2\tau_{1}^{j}+\tau_{11}^{ij}+2\tau_{12}^{jk}+\tau_{112}^{ijk}),$$

that is,

$$2\tau_{11}^{jk}+\tau_{111}^{ijk}=4\tau_{12}^{jk}+2\tau_{112}^{ijk}.$$

It may be verified that the restraints, implied by the relations among the taus, which the estimates must satisfy, are indeed satisfied by those given in table 2. The restraints are:

Case 2: $x_a^*(i.k) = x(i.k)$, $x_a^*(.jk) = x(.jk)$,

$2x^*(ij1) - x_a^*(ij2) - x_a^*(ij3) = 2x(ij1) - x(ij2) - x(ij3)$,

$- x_a^*(ij1) + 2x_a^*(ij2) - x_a^*(ij3) = -x(ij1) + 2x(ij2) - x(ij3)$;

Case 3: $x_b^*(ij.) = x(ij.)$, $x_b^*(i.k) = x(i.k)$,

$x_b^*(1jk) - x_b^*(2jk) = x(1jk) - x(2jk)$;

Case 4a: $x_c^*(ij.) = x(ij.)$, $x_c^*(i.k) = x(i.k)$, $x_c^*(.j2) = x(.j2)$,

$x_c^*(ij2) = x(ij2)$, $x_c^*(1jk) - x_c^*(2jk) = x(1jk) - x(2jk)$;

Case 4b: $x_d^*(ij.) = x(ij.)$, $x_d^*(i.k) = x(i.k)$,

$$x_d^*(1jk) - x_d^*(2jk) = x(1jk) - x(2jk),$$

$$2x_d^*(ij1) + x_d^*(ij2) = 2x(ij1) + x(ij2),$$

$$x_d^*(ij2) + 2x_d^*(ij3) = x(ij2) + 2x(ij3).$$

For convenience we summarize the conditions on the tau parameters associated with each of the estimates.

Case 1: x_2^*, $\tau_{111}^{ijk} = \tau_{112}^{ijk} = 0$,

Case 2: x_e^*, $\tau_{11}^{ij} = \tau_{111}^{ijk} = \tau_{112}^{ijk} = 0$,

Case 2: x_a^*, $3\tau_{11}^{ij} + \tau_{111}^{ijk} + \tau_{112}^{ijk} = 0$,

Case 3: x_f^*, $\tau_{11}^{jk} = \tau_{12}^{jk} = \tau_{111}^{ijk} = \tau_{112}^{ijk} = 0$,

Case 3: x_b^*, $2\tau_{11}^{jk} + \tau_{111}^{ijk} = 0$, $2\tau_{12}^{jk} + \tau_{112}^{ijk} = 0$,

Case 4a: x_g^*, $\tau_{11}^{jk} = \tau_{111}^{ijk} = \tau_{112}^{ijk} = 0$,

Case 4a: x_c^*, $2\tau_{11}^{jk} + \tau_{111}^{ijk} = 0$, τ_{112}^{ijk},

Case 4b: x_h^*, $2\tau_{12}^{jk} = \tau_{11}^{jk} = \tau^{jk}$, $\tau_{111}^{ijk} = \tau_{112}^{ijk} = 0$,

Case 4b: x_d^*, $2\tau_{11}^{jk} + \tau_{111}^{ijk} = 4\tau_{12}^{jk} + 2\tau_{112}^{ijk}$.

As we shall see in section 4, we accept the subhypothesis of case 2 that

$$\tau_{111}^{ijk}=\tau_{112}^{ijk}=\tau_{11}^{ij}=0.$$

It seems reasonable therefore to reexamine the effects of birth-order within the combined model of case 1 and case 2.

Case 5

For the subhypothesis of case 3 within case 1 and case 2, the corresponding estimate $x_m^*(ijk)$ must satisfy

$$\tau_{111}^{ijk}=\tau_{112}^{ijk}=\tau_{11}^{ij}=\tau_{11}^{jk}=\tau_{12}^{jk}=0.$$

The representation for $x_m^*(ijk)$ is given by columns 1-5,7,8 of table 3a. The estimate is explicitly given by

$$x_m^*(ijk) = x(.j.)x(i.k)/n.$$

The numerical values of $x_m^*(ijk)$ are given in table 4. The common logit difference is zero. The parameter values for $x_m^*(ijk)$ are given in table 5.

Case 6a

For the subhypothesis of case 4a within case 1 and case 2, the corresponding estimate $x_n^*(ijk)$ must satisfy

$$\tau_{111}^{ijk}=\tau_{112}^{ijk}=\tau_{11}^{ij}=\tau_{11}^{jk}=0.$$

The representation for $x_n^*(ijk)$ is given by columns 1-5,7,8,10 of table 3a. The values of $x_n^*(ijk)$ are obtained by the iteration (3.6) except that the first cycle is replaced by

$$x^{(3n+1)}(ijk) = (x(.j.)/x^{(3n)}(.j.))x^{(3n)}(ijk) .$$

The numerical values of $x_n^*(ijk)$ are given in table 4. The common logit difference is zero. The parameter values for $x_n^*(ijk)$ are given in table 5.

Case 6b

For the subhypothesis of case 4b within case 1 and case 2, the corresponding estimate $x_p^*(ijk)$ must satisfy

$$\tau_{111}^{ijk}=\tau_{112}^{ijk}=\tau_{11}^{ij}=0,\quad 2\tau_{12}^{jk}=\tau_{11}^{jk}=\tau^{jk}.$$

The representation for $x_p^*(ijk)$ is given by columns 1-5,7-9 of table 3b. The values of $x_p^*(ijk)$ are determined by the iteration (3.8) except that the first cycle is replaced by

$$x^{(3n+1)}(ijk) = (x(.j.)/x^{(3n)}(.j.))x^{(3n)}(ijk).$$

The numerical values of $x_p^*(ijk)$ are given in table 4. The common logit difference is zero. The parameter values for $x_p^*(ijk)$ are given in table 5.

4. The Analysis of Information

To test the various hypotheses and their associated estimates we set up analysis of information tables showing interaction type measures and effect type measures [10], [13]. We recall that interaction type measures compare the original table with an estimated table and test a hypothesis that the values of the tau parameters of the original table satisfy the restraints on the tau parameters specifying the estimated table. This value may also be interpreted as a measure of the

goodness-of-fit of the estimate obtained by fitting the implied relations on the estimated cell values. Effect type measures compare two estimated tables and test a hypothesis that the additional parameters in one of the estimates differ from their zero values in the other estimate. This value may also be interpreted as a measure of the effect of the additional restraints on the estimate in the fitting procedure.

We first compare the estimates within the model of logit additivity with those obtained by using linear contrasts of the logits. In the analysis of information-I we also include the values of the Pearson $\chi^2=\sum\sum\sum(x(ijk)-x_r^*(ijk))^2/x_r^*(ijk)$, for r=a,b,c,d, given by Grizzle [8].

We note that x_a^* and x_e^*, x_b^* and x_f^*, x_c^* and x_g^*, x_d^* and x_h^* do not differ significantly from one another. This is not surprising since from the value of $2I(x:x_2^*)$ we inferred that the values of τ_{111}^{ijk} and τ_{112}^{ijk} for x(ijk) do not differ significantly from $\tau_{111}^{ijk} = \tau_{112}^{ijk} = 0$. However, though the estimates x_a^*,x_b^*,x_c^*,x_d^* and the estimates x_e^*,x_f^*,x_g^*,x_h^* satisfy the linear contrasts for the logits only the latter four satisfy logit additivity. We note that x_e^*,x_f^*,x_g^*,x_h^* are estimated using fewer tau parameters than x_a^*,x_b^*,x_c^*,x_d^* respectively.

Within the model of logit additivity we have analysis of information-II and within case 1 and case 2 analysis of information-III.

Analysis of Information-I

Component due to	Information	χ^2	D.F.
	Case 2		
$\tau^{ij}_{11}=\tau^{ijk}_{111}=\tau^{ijk}_{112}=0$	$2I(x:x^*_e)=3.154$		3
Difference	$2I(x^*_a:x^*_e)=0.654$		2
$3\tau^{ij}_{11}+\tau^{ijk}_{111}+\tau^{ijk}_{112}=0$	$2I(x:x^*_a)=2.500$	2.475	1
	Case 3		
$\tau^{jk}_{11}=\tau^{jk}_{12}=\tau^{ijk}_{111}=\tau^{ijk}_{112}=0$	$2I(x:x^*_f)=28.199$		4
Difference	$2I(x^*_b:x^*_f)=3.724$		2
$2\tau^{jk}_{11}+\tau^{ijk}_{111}=0$, $2\tau^{jk}_{12}+\tau^{ijk}_{112}=0$	$2I(x:x^*_b)=24.475$	24.239	2
	Case 4a		
$\tau^{jk}_{11}=\tau^{ijk}_{111}=\tau^{ijk}_{112}=0$	$2I(x:x^*_g)=25.257$		3
Difference	$2I(x^*_c:x^*_g)=3.932$		2
$2\tau^{jk}_{11}+\tau^{ijk}_{111}=0,\tau^{ijk}_{112}$	$2I(x:x^*_c)=21.325$	22.765	1
	Case 4b		
$2\tau^{jk}_{12}=\tau^{jk}_{11}=\tau^{jk}$, $\tau^{ijk}_{111}=\tau^{ijk}_{112}=0$	$2I(x:x^*_h)=2.131$		3
Difference	$2I(x^*_d:x^*_h)=0.670$		2
$2\tau^{jk}_{11}+\tau^{ijk}_{111}=4\tau^{jk}_{12}+2\tau^{ijk}_{112}$	$2I(x:x^*_d)=1.461$	1.478	1

Analysis of Information-II

Component due to	Information	χ^2	D.F.
	Case 2		
$\tau_{11}^{ij}=\tau_{111}^{ijk}=\tau_{112}^{ijk}=0$	$2I(x:x_e^*)=3.154$		3
Conditional effect of τ_{11}^{ij}	$2I(x_2^*:x_e^*)=2.301$		1
$\tau_{111}^{ijk}=\tau_{112}^{ijk}=0$	$2I(x:x_2^*)=0.853$	0.851	2
	Case 3 and 4a		
$\tau_{11}^{jk}=\tau_{12}^{jk}=\tau_{111}^{ijk}=\tau_{112}^{ijk}=0$ or ($b_1=0$, $b_2=0$)	$2I(x:x_f^*)=28.199$		4
Conditional effect of τ_{12}^{jk} or quadratic component	$2I(x_g^*:x_f^*)=2.942$		1
$\tau_{11}^{jk}=\tau_{111}^{ijk}=\tau_{112}^{ijk}=0$ or ($b_1=0$)	$2I(x:x_g^*)=25.257$		3
Effect of linear component given quadratic component	$2I(x_2^*:x_g^*)=24.404$		1
$\tau_{111}^{ijk}=\tau_{112}^{ijk}=0$	$2I(x:x_2^*)=0.853$		2
	Case 3 and 4b		
	$2I(x:x_f^*)=28.199$		4
Effect of linear component	$2I(x_h^*:x_f^*)=26.068$		1
$2\tau_{12}^{jk}=\tau_{11}^{jk}=\tau^{jk}$, $\tau_{111}^{ijk}=\tau_{112}^{ijk}=0$ or ($b_2=0$)	$2I(x:x_h^*)=2.131$		3
Effect of quadratic component given linear component	$2I(x_2^*:x_h^*)=1.278$		1
	$2I(x:x_2^*)=0.853$		2

Analysis of Information-III

Component due to	Information	D.F.
Case 5 and 6a		
$\tau^{ij}_{11}=\tau^{jk}_{11}=\tau^{jk}_{12}=\tau^{ijk}_{111}=\tau^{ijk}_{112}=0$	$2I(x\ :x^*_m)=29.831$	5
Conditional effect of τ^{jk}_{12} or quadratic component	$2I(x^*_n:x^*_m)=\ 2.904$	1
$\tau^{ij}_{11}=\tau^{jk}_{11}=\tau^{ijk}_{111}=\tau^{ijk}_{112}=0$	$2I(x\ :x^*_n)=26.927$	4
Effect of linear component given quadratic component	$2I(x^*_e:x^*_n)=23.773$	1
$\tau^{ij}_{11}=\tau^{ijk}_{111}=\tau^{ijk}_{112}=0$	$2I(x\ :x^*_e)=\ 3.154$	3
Case 5 and 6b		
$\tau^{ij}_{11}=\tau^{jk}_{11}=\tau^{jk}_{12}=\tau^{ijk}_{111}=\tau^{ijk}_{112}=0$	$2I(x\ :x^*_m)=29.831$	5
Linear component	$2I(x^*_p:x^*_n)=25.399$	1
$2\tau^{jk}_{12}=\tau^{jk}_{11}=\tau^{jk}$, $\tau^{ij}_{11}=\tau^{ijk}_{111}=\tau^{ijk}_{112}=0$	$2I(x\ :x^*_p)=\ 4.432$	4
Quadratic component given linear	$2I(x^*_e:x^*_p)=\ 1.278$	1
$\tau^{ij}_{11}=\tau^{ijk}_{111}=\tau^{ijk}_{112}=0$	$2I(x\ :x^*_e)=\ 3.154$	3

We infer that there is no difference between problems and controls. The effect of birth-order is not homogeneous and the logits may be expressed by a linear regression assuming equal spacing of the effects.

Within the model of logit additivity we may summarize the analysis in the following form.

Source				Information (log-likelihood ratio)	D.F.
Problems vs Controls				3.154	3
Birth-order effects				28.199	4
Linear	26.068	1	Quadratic	2.942	1
Quadratic (conditional)	1.278	1	Linear (conditional)	24.404	1
Interaction	0.853	2	Interaction	0.853	2

Within the model of case 1 and case 2 we may summarize the analysis in the following form

Source				Information (log-likelihood ratio)	D.F.
Birth-order effects and Problems vs Controls				29.831	5
Linear	25.399	1	Quadratic	2.904	1
Quadratic (conditional)	1.278	1	Linear (conditional)	23.773	1
Conditional effect of τ^{ij}_{11} (Problems vs Controls)				2.301	1
Interaction				0.853	2

5. Acknowledgement

This research was partially supported by the Air Force Office of Scientific Research, Office of Aerospace Research, U. S. Air Force Under Grant AFOSR-68-1513. An exchange of correspondence with Professor Grizzle was most helpful. Work with Dr. H. H. Ku in earlier considerations of this problem was most useful. Computer programs were run in the Computer Center of The George Washington University. The capable programming of Mrs. Marian Fisher as well as her interest in the problem and her modifications of a basic program for the analysis of contingency tables by Professor C. T. Ireland contributed to making the theory applicable. I am grateful to Professor J. N. Darroch for the opportunity to participate in the examination of Dr. Douglas Ratcliff's dissertation.

Appendix

We give in this appendix the generalised iterative scaling algorithm as given by Darroch and Ratcliff [5]. Let I be a finite set and let $\underset{\sim}{p} = \{p(i); i\in I, p(i)\geq 0, \sum_{i\in I} p(i)=1\}$ be a probability function on I. Suppose that $\underset{\sim}{p}$ is a member of a family of distributions satisfying the constraints

$$\sum_{i\in I} b_{si}p(i) = k_s, s = 1,2,\ldots d \quad \sum_{i\in I} p(i) = 1 \tag{A.1}$$

where for all s there exist $i\in I$ such that $b_{si}\neq 0$. The constraints in (A.1) may be reformulated into the equivalent canonical form

$$\begin{aligned} &\sum_{i\in I} a_{ri}p(i) = h_r, r = 1,2,\ldots c \\ &a_{ri}\geq 0, \sum_{r=1}^{c} a_{ri} = 1, h_r>0, \sum_{r=1}^{c} h_r = 1 \end{aligned} \tag{A.2}$$

by defining

$$\begin{aligned} &a_{si} = t_s(u_s+b_{si}), \text{ all } i, \\ &h_s = t_s(u_s+k_s), s=1,2,\ldots,d \end{aligned} \tag{A.3}$$

where $u_s\geq 0$, $t_s>0$ are chosen to make

$$a_{si}\geq 0 \text{ and } \sum_{s=1}^{d} a_{si}\leq 1 \text{ for all } i\in I.$$

If $\sum_{s=1}^{d} a_{si}=1$ for all i define c=d, otherwise define c=d+1 and let $a_{ci} = 1 - \sum_{s=1}^{d} a_{si}$, $h_c = 1 - \sum_{s=1}^{d} h_s$.

Now let $\underset{\sim}{\pi}=\{\pi(i),\ i\in I, \pi(i)>0,\ \sum_{i\in I}\pi(i)\leq 1\}$ be a subprobability function on I. The minimum discrimination information estimate $p^*(i)$, $i\in I$, is that member of the family $\underset{\sim}{p}$ satisfying the restraints (A.2) and minimizing

$$I(\underset{\sim}{p};\underset{\sim}{\pi}) = \sum_{i\in I} p(i)\ \ell n\frac{p(i)}{\pi(i)} \tag{A.4}$$

and is given by

$$\ell n\frac{p^*(i)}{\pi(i)} = \sum_{r=1}^{c} a_{ri}\tau_r \tag{A.5}$$

where the τ_r are parameters to be determined so that $p^*(i)$ satisfies the constraints (A.2). The values of $p^*(i)$ may be determined by the convergent iteration

$$p^{(n+1)}(i) = p^{(n)}(i) \prod_{r=1}^{c}\left(\frac{h_r}{h_r^{(n)}}\right)^{a_{ri}}, \quad n=0,1,2,\ldots \tag{A.6}$$

$$p^{(0)}(i) = \pi(i),\ h_r^{(n)} = \sum_{i\in I} a_{ri}p^{(n)}(i).$$

Table 1
Data on number of mothers with previous infant losses (Cochran [4])

No. of mothers with

Birth order k	i		j=1 Losses	j=2 No losses
1	1	Problems	20	82
	2	Controls	10	54
2	1	Problems	26	41
	2	Controls	16	30
3	1	Problems	27	22
	2	Controls	14	23

Table 2
Maximum-likelihood estimates from Berkson [1], Grizzle [8]

i j k	x	Case 1 x_2^*	Case 2 x_a^*	Case 3 x_b^*	Case 4a x_c^*	Case 4b x_d^*
1 1 1	20	20.503	17.812	29.925	27.535	21.191
1 1 2	26	27.213	23.812	22.471	26	23.618
1 1 3	27	25.284	24.812	20.604	19.465	28.191
1 2 1	82	81.497	84.188	72.075	74.465	80.809
1 2 2	41	39.787	43.188	44.529	41	43.382
1 2 3	22	23.716	24.188	28.396	29.535	20.809
2 1 1	10	9.497	12.188	19.925	17.535	11.191
2 1 2	16	14.787	18.188	12.471	16	13.618
2 1 3	14	15.716	16.188	7.604	6.465	15.191
2 2 1	54	54.503	51.812	44.075	46.465	52.809
2 2 2	30	31.213	27.812	33.529	30	32.382
2 2 3	23	21.284	20.812	29.396	30.535	21.809

Table 3a

	1	2	3	4	5	6	7	8	9	10	11	12
i j k	L	τ^{i}_{1}	τ^{j}_{1}	τ^{k}_{1}	τ^{k}_{2}	τ^{ij}_{11}	τ^{ik}_{11}	τ^{ik}_{12}	τ^{jk}_{11}	τ^{jk}_{12}	τ^{ijk}_{111}	τ^{ijk}_{112}
1 1 1	1	1	1	1	0	1	1	0	1	0	1	0
1 1 2	1	1	1	0	1	1	0	1	0	1	0	1
1 1 3	1	1	1	0	0	1	0	0	0	0	0	0
1 2 1	1	1	0	1	0	0	1	0	0	0	0	0
1 2 2	1	1	0	0	1	0	0	1	0	0	0	0
1 2 3	1	1	0	0	0	0	0	0	0	0	0	0
2 1 1	1	0	1	1	0	0	0	0	1	0	0	0
2 1 2	1	0	1	0	1	0	0	0	0	1	0	0
2 1 3	1	0	1	0	0	0	0	0	0	0	0	0
2 2 1	1	0	0	1	0	0	0	0	0	0	0	0
2 2 2	1	0	0	0	1	0	0	0	0	0	0	0
2 2 3	1	0	0	0	0	0	0	0	0	0	0	0
x(ijk)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Case 1 x^*_2	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		
Case 2 x^*_e	✓	✓	✓	✓	✓		✓	✓	✓	✓		
Case 3 x^*_f	✓	✓	✓	✓	✓	✓	✓	✓				
Case 4a x^*_g	✓	✓	✓	✓	✓	✓	✓	✓		✓		
Case 5 x^*_m	✓	✓	✓	✓	✓		✓	✓				
Case 6a x^*_n	✓	✓	✓	✓	✓		✓	✓		✓		

Table 3b

i j k	1 L	2 τ_1^i	3 τ_1^j	4 τ_1^k	5 τ_2^k	6 τ_{11}^{ij}	7 τ_{11}^{ik}	8 τ_{12}^{ik}	9 τ^{jk}
1 1 1	1	1	1	1	0	1	1	0	1
1 1 2	1	1	1	0	1	1	0	1	1/2
1 1 3	1	1	1	0	0	1	0	0	0
1 2 1	1	1	0	1	0	0	1	0	0
1 2 2	1	1	0	0	1	0	0	1	0
1 2 3	1	1	0	0	0	0	0	0	0
2 1 1	1	0	1	1	0	0	0	0	1
2 1 2	1	0	1	0	1	0	0	0	1/2
2 1 3	1	0	1	0	0	0	0	0	0
2 2 1	1	0	0	1	0	0	0	0	0
2 2 2	1	0	0	0	1	0	0	0	0
2 2 3	1	0	0	0	0	0	0	0	0
Case 4b x_h^*	✓	✓	✓	✓	✓	✓	✓	✓	✓
Case 6b x_p^*	✓	✓	✓	✓	✓		✓	✓	✓

Table 4

i j k	x	Case 1 x^*_2	Case 2 x^*_e	Case 3 x^*_f	Case 4a x^*_g	Case 4b x^*_h	Case 5 x^*_m	Case 6a x^*_n	Case 6b x^*_p
1 1 1	20	20.502	18.434	34.156	31.215	22.032	31.578	28.[illegible]	19.844
1 1 2	26	27.212	24.903	22.436	26.811	24.375	20.742	24.903	22.179
1 1 3	27	25.283	23.360	16.408	14.995	26.594	15.170	13.306	24.668
1 2 1	82	81.498	83.566	67.844	70.785	79.968	70.422	71.262	82.156
1 2 2	41	39.788	42.097	44.564	40.191	42.626	46.258	42.097	44.821
1 2 3	22	23.717	25.640	32.592	34.005	22.406	33.830	35.194	24.331
2 1 1	10	9.498	11.566	17.415	15.730	10.256	19.814	18.032	12.452
2 1 2	16	14.788	17.097	12.517	15.189	13.051	14.241	17.097	15.230
2 1 3	14	15.717	17.640	10.068	9.094	16.694	11.455	10.425	18.627
2 2 1	54	54.502	52.434	46.585	48.270	53.744	44.186	45.968	51.548
2 2 2	30	31.212	28.903	33.483	30.809	32.949	31.759	28.903	30.770
2 2 3	23	21.283	19.360	26.932	27.906	20.307	25.545	26.575	18.373

Table 5

Parameter values for various estimates

	Original	Case 1	Case 2	Case 3	Case 4a	Case 4b	Case 5	Case 6a	Case 6b
	x	x_2^*	x_e^*	x_f^*	x_g^*	x_h^*	x_m^*	x_n^*	x_p^*
τ_1^i	-0.0445	0.1083	0.2809	0.1917	0.1977	0.0984	0.2809	0.2809	0.2809
τ_1^j	-0.4964	-0.3031	-0.0931	-0.9839	-1.1212	-0.1959	-0.8020	-0.9358	0.0138
τ_1^k	0.8535	0.9403	0.9963	0.5480	0.5480	0.9733	0.5480	0.5480	1.0316
τ_2^k	0.2657	0.3829	0.4007	0.2177	0.0990	0.4840	0.2177	0.0840	0.5157
τ_{11}^{ij}	0.7012	0.3671	0	0.2977	0.3025	0.3673	0	0	0
τ_{11}^{ik}	0.4622	0.2940	0.1852	0.1852	0.1852	0.2990	0.1852	0.1852	0.1852
τ_{12}^{ik}	0.3568	0.1345	0.0951	0.0952	0.0682	0.1591	0.0951	0.0952	0.0952
τ_{11}^{jk}	-1.1900	-1.4440	-1.4184	0	0	...	0	0	...
τ_{12}^{jk}	-0.1322	-0.4438	-0.4319	0	0.4139	...	0	0.4180	...
τ_{111}^{ijk}	-0.4258	0	0	0	0	0	0	0	0
τ_{112}^{ijk}	-0.5281	0	0	0	0	0	0	0	0
τ^{jk}	...	...	...	...	...	-1.4605	...	...	-1.4344

Table 6

	111	112	113	121	122	123	211	212	213	221	222	223
$a_1(ijk)$	1	1/2	0	0	0	0	1	1/2	0	0	0	0
$a_2(ijk)$	0	0	0	1	1/2	0	0	0	0	1	1/2	0
$a_3(ijk)$	0	1/2	1	0	1/2	1	0	1/2	1	0	1/2	1

$$h_1 = x(111)+\tfrac{1}{2}x(112)+x(211)+\tfrac{1}{2}x(212)=x(.11)+\tfrac{1}{2}x(.12)$$

$$h_2 = x(121)+\tfrac{1}{2}x(122)+x(221)+\tfrac{1}{2}x(222)=x(.21)+\tfrac{1}{2}x(.22)$$

$$h_3 = \tfrac{1}{2}x(112)+x(113)+\tfrac{1}{2}x(122)+x(123)+\tfrac{1}{2}x(212)+x(213)+\tfrac{1}{2}x(222)$$

$$+x(223)=x(..3)+\tfrac{1}{2}x(..2)$$

References

[1] Berkson, J. [1968]. Application of minimum logit χ^2 estimate to a problem of Grizzle with a notation on the problem of no interaction. Biometrics 24, 75-96.

[2] Bishop, Y. M. M. [1971]. Effects of collapsing multi-dimensional contingency tables. Biometrics 27, 545-62.

[3] Campbell, L. L. [1970]. Equivalence of Gauss's principle and minimum discrimination information estimation of probabilities. Annals of Mathematical Statistics 41, 1011-15.

[4] Cochran, W. G. [1954]. Some methods for strengthening the common Chi-square tests. Biometrics 10, 417-51.

[5] Darroch, J. N. and Ratcliff, D. [1972]. Generalised iterative scaling and maximum likelihood. The Annals of Mathematical Statistics (to appear).

[6] Dempster, A. P. [1971]. An overview of multivariate data analysis. Journal of Multivariate Analysis 1, 316-47.

[7] Gokhale, D. V. [1971]. An iterative procedure for analysing log-linear models. Biometrics 27, 681-87.

[8] Grizzle, J. E. [1961]. A new method of testing hypotheses and estimating parameters for the logistic model. Biometrics 17, 372-385.

[9] Kendall, M. G. and Stuart, A. [1961]. The Advanced Theory of Statistics, Vol. 2, Charles Griffin & Company, Ltd., London.

[10] Ku, H. H., Varner, R. N., Kullback, S. [1971]. On the analysis of multidimensional contingency tables. Journal of the American Statistical Association 66, 55-64.

[11] Kullback, S. [1959]. Information Theory and Statistics, Wiley, N. Y., 1968, Dover Publications, N. Y.

[12] Kullback, S. [1970]. Various applications of minimum discrimination information estimation, particularly to problems of contingency table analysis. Proceedings of the Meeting on Information Measures. Kitchener-Waterloo, Ontario, Canada, April 10-14, 1970. I-33-66.

[13] Kullback, S. [1970]. Minimum discrimination information estimation and application. Invited paper presented to SIXTEENTH CONFERENCE ON THE DESIGN OF EXPERIMENTS IN ARMY RESEARCH, DEVELOPMENT AND TESTING, U.S.A. LOGISTICS MANAGEMENT CENTER, FT. LEE, VA. 21 OCTOBER 1970. ARO-D Report 71-3, 1-38, Proceedings of the conference.

[14] Plackett, R. L. [1969]. Multidimensional contingency tables: A survey of models and methods. Bulletin of the International Statistical Institute, Book 1, 43, 133-42.

[15] Snedecor, G. W. and Cochran, W. G. [1967]. Statistical Methods, The Iowa State University Press, Ames, Iowa.